Version with markings to show changes made

Oligonucleotide D1024 (GTN TG(T/C) GA(T/C) GGN TT(T/C) CA(T/C) GTN GG) (Seq ID NO 1) has a degeneracy of 1024. The sequence in which dK is substituted for the A/G degenerate position, dP for the C/T degenerate position and dP/dK for all four nucleotides is given by GT(dP/dK) TG(dP) GA(dP) GG(dP/dK) TT(dP) CA(dP) GT(dP/dK) GG (Seq ID NO 12). The sequences represented by oligonucleotide D1024-PK, which has a degeneracy of 8, are depicted in Table 3.

In another embodiment of the present invention, the present invention may be used to preferentially isolate cDNA molecules that contain the 5' terminus including the translation initiation codon. This is accomplished by developing degenerate oligonucleotide to the Kozak sequence which includes the translation initiation codon and extends 5' approximately 13 nucleotides (Kozak, M, *Nucleic Acids Res. 8*:125-32 (1987); Kozak, M, *J. Biol. Chem 266*:19867-70 (1991)). The consensus sequence for initiation of translation by eukaryotic ribosomes is [GCC GCC A^{-3} /GCC A $^{1}UGG^{4}$] GCC GCC $(A/G)^{-3}$CC A^{1}UG G^{4} (SEQ ID NO 11), Kozak, M, *Nucleic Acids Res. 8*:125-32 (1987); Kozak, M, *J. Biol. Chem 266*:19867-70 (1991), herein incorporated by reference; Sambrook *et al.*, 16.16, In *Molecular Cloning, a Laboratory Manual*, Cold Spring Harbor Press (1989), herein incorporated by reference. Two approaches can be attempted to enrich for the presence of the 5' terminus including the translation start codon. In the first, the degenerate Kozak oligonucleotide [prbe] probe can be used to enrich by GeneTrapper for 5' sequences followed by the use of a gene-specific

GeneTrapper probe. Alternatively, a gene-specific GeneTrapper probe can be applied to a phagemid cDNA library using GeneTrapper followed by the use of a degenerate Kozak oligonucleotide probe. In both cases, the percentage of clones that contain the 5' terminus including the translation initiation codon [shoule] should be enriched. This method will be especially useful for clones derived from longer mRNAs (i.e., greater than 5 Kb).

The sequence listing has been replaced with a substitute sequence listing.